REMARKS

The Applicant thanks the Examiner for the telephonic interview with the Applicant's representative, J. Robin Rohlicek, on February 10, 2005, at which time *Kuhn* (US Pat. 6,029,132) was discussed and the Examiner agreed to withdraw the final rejection of the claims over *Kuhn* and over *Kuhn* in view of *Mohri* (US Pat. 6,243,679). In the present office action, claims 1-10 and 18 now stand rejected as anticipated by *Hunt (IEEE 1996),* which was cited by the Applicant prior to the first office action, and claims 4 and 11-17 stand rejected as obvious over *Hunt* alone (claim 4), or in view of *Beutenagel* (Eurospeech 1999) (claims 11-16), also cited by the Applicant, or in view of *Mohri* (claim 17).

Before specifically addressing the rejections of the claims, the *Hunt, Beutenagel,* and *Mohri* references, which may be relevant to the following discussion, are briefly presented. This discussion of the references should not be construed to be a statement regarding the scope of the claims.

The Applicant recognizes that *Hunt* uses a state transition network for the purpose of determining a sequence of database units $u_1^n = (u_1,\ldots,u_t)$ (each corresponding to a different waveform segment) for synthesizing a target specification $t_1^n = (t_1,\ldots,t_t)$. Each element of the target specification, t_i, is synthesized by a database unit, u_i, with the same identify and that target element. For example, if phonemes are used, the target specification corresponds to a phonetic pronunciation, each target phoneme is synthesized by a database unit with the same phonetic identity as the target phoneme. The sequence of database units for a particular target specification is found using a graph-based search, in particular, using a hidden-Markov model (HMM) technique (see page 2, column 1, lines 19-31).

The graph-based search involves computing a "concatenation cost" between pairs of database units. For any pair of database units, u_{i-1},u_i, the concatenation cost $C^c(u_{i-1},u_i)$ depends on features of the waveform segments associated with the database units, for example "cepstral distance at the point of concatenation and the absolute difference in log power and

pitch." (see page 2, column 2, lines 33-41). Note that the concatenation costs in *Hunt* are not represented in the network used for the search.

Beutnagel provides an improvement to the *Hunt* approach. Beutnagel addresses the problem that computation of "join costs" (i.e. *Hunt's* "concatenation costs") "can be quite expensive to compute" (abstract, lines 6-7), and describes an approach in which a subset of all possible concatenation costs are precomputed and cached.

Mohri presents an approach to use of Finite State Transducers (FSTs) for speech recognition (i.e. speech waveform to text representation). *Mohri* asserts that " These algorithms also apply to text-to-speech synthesis" (col. 24, lines 36-37). However, *Mohri* does not provide any specific suggestion regarding how such algorithms might be applied. In *Mohri*, FST are labeled with and accept units such as phonetic units, and words. (see col. 4, lines 11-18). Transitions, for example between phonetic labels, are not associated with transition labels.

Without intending to limit the claims to any particular embodiment, an embodiment described in the present application introduces a cost of concatenating segments without necessarily having to consider specific characteristics of the segments, such as signal characteristics. (page 4, lines 3-5). For example, concatenation costs are based on symbolic labeling of the segments. (page 6, lines 14-16). In the embodiment illustrated in FIG. 5, these concatenation costs are embodied on the constraint kernel (520) that links segments that come from different source utterances and includes elements associated with transition labels. (see generally beginning on page 11, line 15).

Turning now to the claims, claim 1 has been amended to make clear that a path through the graph "determin[es] a numerical a numerical score that characterizes a quality of a concatenation of the sequence of segments based on quantities characterizing elements of the graph." This is in contrast to *Hunt* (and to *Beutnagel)* which teaches such a score being determined not only from the path but from characteristics of the pairs of concatenated database units themselves (e.g., from a cepstral distance, which is a measure of spectral discontinuity, at the point of concatenation of a pair of units). Because all pairs of database units are not represented, for example, as transitions in the prior art graphs, their concatenation costs cannot

be considered to be "quantities characterizing elements of the graph." In contrast, in an embodiment disclosed in the present applications such quantities are supported by the concatenation costs that characterize links in the constraint kernel (520). The prior art, for instance *Beutnagel,* recognizes that "the search space of possible joins is large" and therefore not easily represented in a graph. *Beutnagel,* rather than representing the concatenation costs by quantities characterizing elements of a graph, uses a cache to enable precomputation and fast lookup of quantities as they are needed. Likely because the search space of possible pairs of potentially concatenated units is large (i.e., growing generally as the square of the number of units), none of the cited references discloses or suggests using quantities characterizing such a large number of elements (i.e., links) of a graph.

Claim 1 has also been amended to require that each path through the graph identifies a sequence of unit labels and transition labels, and that the target utterance is represented by until labels and transition labels. *Hunt* does not disclose or suggest use of such transition labels.

Claim 11 has been rewritten in independent form to include the limitations of that claim as previously pending. Claim 11 stands rejected as obvious over *Hunt* in view of *Beutnagel.* The office action recognizes that *Hunt* the limitations recited in dependent claim 11 as previously pending and relies on *Beutnagel* to provide what is missing. In particular, the office action takes the position that *Beutnagel's* pre-tabulation of concatenation costs in a cache discloses the recited "second part" of the graph that (a) "encodes allowable transitions between segments of different source utterances" and (b) "encodes a transition score of each of those transitions." Although *Beutnagel* may tabulate in a cache the transition scores between segments of different source utterances, this cache is not part of a graph – *Beutnagel* uses a data structure that is separate from the graph and cannot be considered to be in any way included in the graph as required by the claim.

Claim 14 has also been amended to be an independent claim. As with the rejection of claim 11, the office action takes the position that *Beutnagel's* pre-tabulation of concatenation costs in a cache discloses the recited "second part" of the graph. The applicant disagrees for the reason set forth above for claim 11.

Claim 14 as amended also requires that the first part of the graph encodes a sequence of unit labels and transition labels for each of the source utterances. Neither *Hunt* nor *Beutnagel's* disclose or suggest a first part that encodes a sequence of labels for each of the source utterance. Once *Hunt's* network is constructed, there is no encoding of the sequence for any particular source utterance - all the segments are combined irrespective of their source utterances.

In addition, neither *Hunt* nor *Beutnagel* disclose or suggest a first part that encodes a sequence of unit labels and transition labels for each of the source utterances. The references do not use transition labels and there is no reason to suggest they would have been motivated to introduce transition labels into their approaches.

Dependent claim 17 stands rejected as obvious over *Hunt* in view of *Mohri*. The office action asserts that *Mohri* discloses an FST that accepts unit labels and transition labels. However, nowhere does *Mohri* in fact disclose an FST that accepts transition labels as required by the claim.

New dependent claims 19-20 have been added dependent on claim 14. These claims recite features that generally relate to limitations recited in pending claim 17, which depends on claim 1.

It is believed that all of the pending claims have been addressed. However, the absence of a reply to a specific rejection, issue or comment does not signify agreement with or concession of that rejection, issue or comment. In addition, because the arguments made above may not be exhaustive, there may be reasons for patentability of any or all pending claims (or other claims) that have not been expressed. Finally, nothing in this paper should be construed as an intent to concede any issue with regard to any claim, except as specifically stated in this paper, and the amendment of any claim does not necessarily signify concession of unpatentability of the claim prior to its amendment.

Please apply the $510.00 Petition for Extension of Time fee and any other charges or credits to deposit account 06-1050.

Attorney's Docket No.: 01997-294001 / Case No. 8973

Respectfully submitted,

Date: Jan. 25, 2006

J. Robin Rohlicek
Reg. No. 43,349

Fish & Richardson P.C.
225 Franklin Street
Boston, MA 02110
Telephone: (617) 542-5070
Facsimile: (617) 542-8906

21252440.doc